

S. Liu

RUSH



ENTERED

1653

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/707,167

DATE: 08/15/2002

TIME: 15:10:19

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw

```

3 <110> APPLICANT: Moloney, Maurice
4      Boothe, Joseph
5      Rooijen, Gijs Van
7 <120> TITLE OF INVENTION: Oil Bodies and Associated Proteins as Affinity Matrices
9 <130> FILE REFERENCE: 9369-161
11 <140> CURRENT APPLICATION NUMBER: US 09/707,167
12 <141> CURRENT FILING DATE: 2000-11-07
14 <150> PRIOR APPLICATION NUMBER: US 09/319,275
15 <151> PRIOR FILING DATE: 1999-08-27
17 <160> NUMBER OF SEQ ID NOS: 14
19 <170> SOFTWARE: PatentIn version 3.1
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 522
23 <212> TYPE: DNA
24 <213> ORGANISM: Arabidopsis Thaliana
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (1)..(522)
29 <223> OTHER INFORMATION:
32 <400> SEQUENCE: 1
33 atg gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga gac      48
34 Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
35 1          5          10          15
37 cag tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc gga      96
38 Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
39          20          25          30
41 cga gga tct gac tac tcc aag tct agg cag att gct aaa gct gca act      144
42 Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
43          35          40          45
45 gct gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt      192
46 Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
47          50          55          60
49 gtt gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt atc      240
50 Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
51 65          70          75          80
53 ttc agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc      288
54 Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
55          85          90          95
57 acc ggt ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc gtt      336
58 Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
59          100          105          110
61 ttc tct tgg att tac aag tac gca acg gga gag cac cca cag gga tca      384
62 Phe Ser Trp Ile Tyr Lys Tyr Ala Thr Gly Glu His Pro Gln Gly Ser

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/707,167

DATE: 08/15/2002

TIME: 15:10:19

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw

```

63          115          120          125
65 gac aag ttg gac agt gca agg atg aag ttg gga agc aaa gct cag gat      432
66 Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
67          130          135          140
69 ctg aaa gac aga gct cag tac tac gga cag caa cat act ggt ggg gaa      480
70 Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
71 145          150          155          160
73 cat gac cgt gac cgt act cgt ggt ggc cag cac act act taa      522
74 His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr
75          165          170
78 <210> SEQ ID NO: 2
79 <211> LENGTH: 173
80 <212> TYPE: PRT
81 <213> ORGANISM: Arabidopsis Thaliana
83 <400> SEQUENCE: 2
85 Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
86 1          5          10          15
89 Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
90          20          25          30
93 Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
94          35          40          45
97 Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
98          50          55          60
101 Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
102 65          70          75          80
105 Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
106          85          90          95
109 Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
110          100          105          110
113 Phe Ser Trp Ile Tyr Lys Tyr Ala Thr Gly Glu His Pro Gln Gly Ser
114          115          120          125
117 Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
118          130          135          140
121 Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
122 145          150          155          160
125 His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr
126          165          170
129 <210> SEQ ID NO: 3
130 <211> LENGTH: 2115
131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: Oleosin - Hirudin Fusion
137 <220> FEATURE:
138 <221> NAME/KEY: CDS
139 <222> LOCATION: (862)..(1215)
140 <223> OTHER INFORMATION:
143 <220> FEATURE:
144 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING

DATE: 08/15/2002

PATENT APPLICATION: US/09/707,167

TIME: 15:10:19

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw

```

145 <222> LOCATION: (1456)..(1833)
146 <223> OTHER INFORMATION:
149 <400> SEQUENCE: 3
150 ctatacccaa cctcggctctt ggtcacacca ggaactctct ggtaagctag ctccactccc      60
152 cagaaacaac cggcgccaaa ttgccggaat tgctgacctg aagacggaac atcatcgctcg      120
154 ggtccttggg cgattgcggc ggaagatggg tcagcttggg cttgaggacg agaccggaat      180
156 cgagtctgtt gaaaggttgt tcattgggat ttgtatacgg agattggctg tcgagagggt      240
158 tgagggaaag gacaaatggg tttggctctg gagaaagaga gtgcggcttt agagagagaa      300
160 ttgagagggt tagagagaga tgcggcggcg atgacgggag gagagacgac gaggacctgc      360
162 attatcaaag cagtgcgtg gtgaaatttg gaacttttaa gaggcagata gatttattat      420
164 ttgtatccat tttcttcatt gttctagaat gtgcgggaac aaattttaaa actaaatcct      480
166 aaatttttct aattttgttg ccaatagtgg atatgtgggc cgtatagaag gaatctattg      540
168 aaggcccaaa cccatactga cgagcccaaa ggttcgtttt gcgttttatg tttcggttcg      600
170 atgccaacgc cacattctga gctaggcaaa aaacaaacgt gtctttgaat agactcctct      660
172 cgttaacaca tgcagcggct gcatggtgac gccattaaca cgtggcctac aattgcatga      720
174 tgtctccatt gacacgtgac ttctcgtctc cttctttaat atatctaaca aacactccta      780
176 cctcttccaa aatatataca catctttttg atcaatctct cattcaaat ctcattctct      840
178 ctagtaaaaa agaacaaaaa a atg gcg gat aca gct aga gga acc cat cac      891
179                               Met Ala Asp Thr Ala Arg Gly Thr His His
180                               1                               5                               10
182 gat atc atc ggc aga gac cag tac ccg atg atg ggc cga gac cga gac      939
183 Asp Ile Ile Gly Arg Asp Gln Tyr Pro Met Met Gly Arg Asp Arg Asp
184                               15                               20                               25
186 cag tac cag atg tcc gga cga gga tct gac tac tcc aag tct agg cag      987
187 Gln Tyr Gln Met Ser Gly Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln
188                               30                               35                               40
190 att gct aaa gct gca act gct gtc aca gct ggt ggt tcc ctc ctt gtt      1035
191 Ile Ala Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val
192                               45                               50                               55
194 ctc tcc agc ctt acc ctt gtt gga act gtc ata gct ttg act gtt gca      1083
195 Leu Ser Ser Leu Thr Leu Val Gly Thr Val Ile Ala Leu Thr Val Ala
196                               60                               65                               70
198 aca cct ctg ctc gtt atc ttc agc cca atc ctt gtc ccg gct ctc atc      1131
199 Thr Pro Leu Leu Val Ile Phe Ser Pro Ile Leu Val Pro Ala Leu Ile
200 75                               80                               85                               90
202 aca gtt gca ctc ctc atc acc ggt ttt ctt tcc tct gga ggg ttt ggc      1179
203 Thr Val Ala Leu Leu Ile Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly
204                               95                               100                               105
206 att gcc gct ata acc gtt ttc tct tgg att tac aag taagcacaca      1225
207 Ile Ala Ala Ile Thr Val Phe Ser Trp Ile Tyr Lys
208                               110                               115
210 tttatcatct tacttcataa ttttgtgcaa tatgtgcatg catgtgttga gccagtagct      1285
212 ttggatcaat ttttttggtc gaataacaaa tgtaacaata agaaattgca aattctaggg      1345
214 aacatttggg taactaaata cgaaatttga cctagctagc ttgaatgtgt ctgtgtatat      1405
216 catctatata ggtaaaatgc ttggtatgat acctattgat tgtgaatagg tac gca      1461
217                               Tyr Ala
218                               120
220 acg gga gag cac cca cag gga tca gac aag ttg gac agt gca agg atg      1509
221 Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala Arg Met

```

RAW SEQUENCE LISTING

DATE: 08/15/2002

PATENT APPLICATION: US/09/707,167

TIME: 15:10:19

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw

```

222          125          130          135
224 aag ttg gga agc aaa gct cag gat ctg aaa gac aga gct cag tac tac      1557
225 Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln Tyr Tyr
226          140          145          150
228 gga cag caa cat act ggt tgg gaa cat gac cgt gac cgt act cgt ggt      1605
229 Gly Gln Gln His Thr Gly Trp Glu His Asp Arg Asp Arg Thr Arg Gly
230          155          160          165
232 ggc cag cac act act gcg atc gaa ggg aga atc act tac act gac tgt      1653
233 Gly Gln His Thr Thr Ala Ile Glu Gly Arg Ile Thr Tyr Thr Asp Cys
234          170          175          180
236 act gaa tct gga cag aac ctc tgt ctc tgt gaa gga tct aac gtt tgt      1701
237 Thr Glu Ser Gly Gln Asn Leu Cys Leu Cys Glu Gly Ser Asn Val Cys
238 185          190          195          200
240 gga aag gga aac aag tgt atc ctc gga tct aac gga aag gga aac cag      1749
241 Gly Lys Gly Asn Lys Cys Ile Leu Gly Ser Asn Gly Lys Gly Asn Gln
242          205          210          215
244 tgt gtt act gga gaa gga act cca aac cca gaa tct cac aac aac gga      1797
245 Cys Val Thr Gly Glu Gly Thr Pro Asn Pro Glu Ser His Asn Asn Gly
246          220          225          230
248 gac ttc gaa gaa atc cct gaa gaa tac ctc cag taa gtcgactcta      1843
249 Asp Phe Glu Glu Ile Pro Glu Glu Tyr Leu Gln
250          235          240
252 gacggatctc ccgatcggtc aaacatttgg caataaagtt tcttaagatt gaatcctgtt      1903
254 gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt      1963
256 aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta      2023
258 tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc      2083
260 gcggtgtcat ctatgttact agatcggaat tc      2115
263 <210> SEQ ID NO: 4
264 <211> LENGTH: 118
265 <212> TYPE: PRT
266 <213> ORGANISM: Artificial Sequence
268 <220> FEATURE:
269 <223> OTHER INFORMATION: Oleosin - Hirudin Fusion
271 <400> SEQUENCE: 4
273 Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
274 1          5          10          15
277 Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
278          20          25          30
281 Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
282          35          40          45
285 Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
286          50          55          60
289 Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
290 65          70          75          80
293 Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
294          85          90          95
297 Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
298          100          105          110
301 Phe Ser Trp Ile Tyr Lys

```

RAW SEQUENCE LISTING

DATE: 08/15/2002

PATENT APPLICATION: US/09/707,167

TIME: 15:10:19

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw

```

302      115
305 <210> SEQ ID NO: 5
306 <211> LENGTH: 125
307 <212> TYPE: PRT
308 <213> ORGANISM: Artificial Sequence
310 <220> FEATURE:
311 <223> OTHER INFORMATION: Oleosin - Hirudin Fusion
313 <400> SEQUENCE: 5
315 Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
316 1      5      10      15
319 Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
320      20      25      30
323 Tyr Tyr Gly Gln Gln His Thr Gly Trp Glu His Asp Arg Asp Arg Thr
324      35      40      45
327 Arg Gly Gly Gln His Thr Thr Ala Ile Glu Gly Arg Ile Thr Tyr Thr
328      50      55      60
331 Asp Cys Thr Glu Ser Gly Gln Asn Leu Cys Leu Cys Glu Gly Ser Asn
332 65      70      75      80
335 Val Cys Gly Lys Gly Asn Lys Cys Ile Leu Gly Ser Asn Gly Lys Gly
336      85      90      95
339 Asn Gln Cys Val Thr Gly Glu Gly Thr Pro Asn Pro Glu Ser His Asn
340      100     105     110
343 Asn Gly Asp Phe Glu Glu Ile Pro Glu Glu Tyr Leu Gln
344      115     120     125
347 <210> SEQ ID NO: 6
348 <211> LENGTH: 2366
349 <212> TYPE: DNA
350 <213> ORGANISM: Artificial Sequence
352 <220> FEATURE:
353 <223> OTHER INFORMATION: Oleosin - Metallothionein Fusion
355 <220> FEATURE:
356 <221> NAME/KEY: CDS
357 <222> LOCATION: (1092)..(1856)
358 <223> OTHER INFORMATION:
361 <400> SEQUENCE: 6
362 gagctcaaat acgatctgat actgataacg tctagatttt taggggttaa gcaatcaatc      60
364 acctgacgat tcaaggtggt tggatcatga cgattccaga aaacatcaag caagctctca      120
366 aagctacact ctttgggatc atactgaact ctaacaacct cgttatgtcc cgtagtgcca      180
368 gtacagacat cctcgtaact cggattatgc acgatgccat ggctataccc aacctcggtc      240
370 ttggtcacac caggaactct ctggtaaagt agctccactc cccagaaaaca accggcgcca      300
372 aattgccgga attgctgacc tgaagacgga acatcatcgt cgggtccttg ggcgattgcg      360
374 gcggaagatg ggtcagcttg ggcttgagga cgagaccgga atcgagtctg ttgaaagggt      420
376 gttcattggg atttgataac ggagattggt cgtcgagagg tttgagggaa aggacaaatg      480
378 ggtttggctc tggagaaaaga gagtgcggct ttagagagag aattgagagg tttagagaga      540
380 gatgcggcgg cgatgacggg aggagagacg acgaggacct gcattatcaa agcagtgcg      600
382 tggtgaaatt tggaactttt aagaggcaga tagattttatt atttgtatcc attttcttca      660
384 ttgttctaga atgtcgcgga acaaatttta aaactaaatc ctaaattttt ctaattttgt      720
386 tgccaatagt ggatatgtgg gccgtataga aggaatctat tgaaggccca aaccatact      780
388 gacgagccca aagggttcgtt ttgcgtttta tgtttcgggt cgatgccaac gccacattct      840

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/707,167

DATE: 08/15/2002

TIME: 15:10:20

Input Set : A:\SEQUENCE.txt

Output Set: N:\CRF4\08152002\I707167.raw